DOCUMENT RESUME

ED 311 083 TM 013 927

TITLE	Recommendations for State-by-State Assessment of Student Achievement.
INSTITUTION	Council of Chief State School Officers, Washington, DC. State Education Assessment Center.
SPONS AGENCY	Center for Statistics (OERI/ED), Washington, DC.
PUB DATE	17 Oct 86
CONTRACT	300-85-0146
NOTE	10p.; For related documents, see TM 013 919-920, TM 013 922-923, TM 013 925, TM 013 928-929.
AVAILABLE FROM	Council of Chief State School Officers, State Education Assessment Center, 400 North Capitol Street, N.W., Washington, DC 20001 (free).
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	*Academic Achievement; *Achievement Rating; Basic Skills; *Educational Assessment; *Evaluation Methods; *Program Development; State Programs; Statewide Planning
IDENTIFIERS	Council of Chief State School Officers; *State by State Assessment; *State Education Assessment Center

ABSTRACT

Recommendations are presented for the state-by-state assessment of student achievement as part of the development of educational indicators to be compared across states. The State Education Assessment Center of the Council of Chief State School Officers solicited comments from the states about developing the assessment program. Issues were addressed in a meeting in March 1986 that resulted in the following provisions for state-by-state assessment: (1) that states be convened to establish the broadest possible statements on content and priorities in each subject upon which they can agree; (2) that assessment results be reported in simple, global format in each subject at each grade level; (3) that assessment of reading, writing, and English be alternated with mathematics, so that basic skills are assessed every other year, with science and social studies assessed every fourth year; and (4) that the Center establish explicit guidelines for all significant procedures in the assessment program. The plan adopted in 1985 had established that assessment activities would be pursued in conjunction with the National Assessment of Educational Progress (NAEP). If research shows that NAEP items and exercises are not suitable in the context of these recommendations, assessment activities are to be developed independently with the states. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

COUNCIL OF CHIEF STATE SCHOOL OFFICERS
State Education Assessment Center

Education Data Improvement Project

U.S. DEPARTMENT OF EDUCATION
Center for Statistics

ED311083

RECOMMENDATIONS FOR STATE-BY-STATE ASSESSMENT OF STUDENT ACHIEVEMENT

Presented to:
Committee on Coordinating Educational Information and Research
Council of Chief State School Officers

Presented by:
Ramsay Selden, Director
State Education Assessment Center

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

October 17, 1986

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY
W. F. Pierce
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Prepared by:
State Education Assessment Center
Council of Chief State School Officers
377 Hall of the States
400 North Capitol St. NW
Washington, DC 20001
(202) 624-7700

TM013927

BEST COPY AVAILABLE

Recommendations:

State-by-State Assessment of Student Achievement

Based on planning efforts with the states over the past year, the following provisions are recommended for state-by-state assessment of student achievement.

1. Establishing Subject Matter to Be Assessed. The CEIR Committee plans adopted by CCSSO in 1985 called for assessment in reading, mathematics, science, social studies, and English literature, writing, and language arts. The plans did not indicate how the content and skills to be assessed in each subject were to be determined. Based on discussions with the states, it is recommended that the states be convened to establish the broadest possible or optimal statements on content and priorities in each subject that they can agree upon, as opposed to gearing assessment to a minimal set of skills that are pursued in common among the states.

2. Determining the level of information to be reported from the program. The 1985 CEIR Committee report called for appropriate scales to be used in reporting student achievement information, based on the specification of subject matter and the item pool that is used. It is recommended that state-by-state results from the assessment program be reported in a simple, global format for performance in each subject at each grade level. Additional data, such as performance of major sub-areas in a subject (like comprehension in reading), will reported only as optional or supplementary information. The scale to be used in reporting information will be appropriate to the item pool that is ultimately used in the program.

3. Establishing the schedule and cycle for assessing subjects: The 1985 CEIR Committee plan established that subjects were to be assessed on a cyclical basis, but did not specify how often or in what order. It is recommended that assessment of reading, writing, and English be alternated with mathematics, so that these basic skills are assessed every other year. Science and social studies will be alternated in combination with mathematics assessments, so each of them will be assessed every four years. Under the schedule set by CEIR for implementating the assessment program, reading, writing, and English will be assessed in the 1988-89 school year, mathematics and science in 1989-90, reading, writing, and English again in 1990-91, and mathematics and social studies in 1991-92. Each assessment will be done at the elementary, intermediate, and secondary school levels.

4. Ensuring that assessment procedures are coordinated and uniform across states. The CEIR committee recommended that the CCSSO Assessment Center coordinate administration of the assessment program, but did

not specify how such administration was to be structured. It is recommended that the Center establish explicit guidelines covering all significant procedures in the assessment program, and that the Center monitor conformity with these guidelines as assessment is done in the states. In addition, it is recommended that sampling, scoring, and data processing for the program be centralized and supported by resources pooled from the states and administered by the Assessment Center or another suitable organization.

The CEIR Committee plan adopted by the Council in 1985 established that these activities be pursued first in conjunction with the National Assessment of Educational Progress, and that they would be developed independently with the states if NAEP is not found to be appropriate. With the approval of these recommendations by the CEIR Committee, they will be conveyed to the U.S. Department of Education Center for Statistics as design specifications to use as a basis for negotiating the content of the next period of the National Assessment project to be extended in 1988.

If NAEP items and exercises are found to be appropriate and are used in the assessment, NAEP procedures for test-administration will be adapted and followed, NAEP reporting scales will be considered and used if found appropriate, students will be assessed in age and grade levels consistent with NAEP (grades 3, 7, and 11), and a sample of students which permit reporting of results by sex, ethnic group, and type of community will be assessed each year. This would entail sampling and testing about 2,000 students per state per grade level, or a total of about 6,000 students per state. Finally, student testing time in each subject would be about 50 minutes, based on past experience with NAEP. It is suggested that consideration be given to testing a single sample of students in both subjects on different occasions, in those years when more than one subject is assessed, in order to maximize economy and efficiency.

Background

In November, 1985, the Council of Chief State School Officers adopted a plan developed by the Committee on Coordinating Educational Information and Research for developing educational indicators to be compared across states. These indicators included a measure of student achievement. The CEIR Committee plans provided for the development of a common measure of student achievement to be collected across states beginning with the 1988-89 school year. In addition, the plans called for measuring student achievement on a cyclical basis in the subjects of reading, mathematics, science, social studies, and English, at the elementary, intermediate, and secondary school levels, drawing on a common test-item pool. The plan called for considering first the item pool developed in the National Assessment of Educational Progress for this purpose.

With these parameters as a base, work began last winter in the CCSSO State Education Assessment Center to proceed with the development of a state-by-state program for assessing student achievement. Comment was sought from states on three major issues around which consensus had to be established in order to proceed with development of the program: how would states get together on the specification of subject matter to be assessed? What kind of information should be provided from the assessment program, including the cycle for the assessment program; that is, on what schedule should the various subject areas be assessed? And, how could uniform collection of data in the states be ensured; that is, what provisions did states want, in order to ensure standardized data-collection procedures across states?

These issues were addressed in a meeting last March of testing and instructional program staff from thirty-five states. Each issue was presented, options and implications were discussed, and either a preliminary consensus or major alternatives were developed. These discussions and recommendations were conveyed to states in a white paper, and the responses of individual states were sought to the issues presented in the white paper. Responses were received from twenty states and have been compiled and analyzed, permitting development of recommendations for the student achievement assessment program.

The remainder of this paper presents those conclusions, organized around the three major issues and presenting the position of states on the design issues.

Specifying Subject Matter

Development of any educational assessment program must begin with specification of the subject matter to be assessed. This must be done to allow selection or development of a test that corresponds validly to the content and emphases of the instructional program being assessed.

This process is critical when developing an assessment program at the state or national levels. Because of the diversity that exists among the instructional programs operated by states, local school districts, schools, and teachers, subject matter must not only be defined; consensus or commonality among differing programs must be established.

Three options were developed for arriving at such commonality. First, the instructional objectives or curricular frameworks of those states that have developed them could be examined, so that knowledge and skills that were common across all of them would be identified. This was referred to as the "least common denominator" approach, because it would result in specification only of subject matter that was common across all states, and would include nothing that was not addressed by all states that have established curricular objectives or priorities. As a second approach, states could be convened to establish the maximal set of subject-matter and priorities that they can agree upon, even though some of the subject matter or the priorities would differ somewhat from the states' own objectives and priorities. Here, it was felt that a maximal consensus could be established, based on the collective opinion of states as to what was educationally important, even though it might go beyond what some had articulated as curricular policy. Finally, a mixed option was discussed, in which a common core would be established to be assessed in all states, and in which instructional components in addition to the core could be assessed at the discretion of the individual state.

The response from states overwhelmingly favored the "maximal consensus" approach. Only one state favored the "least common denominator," and one favored the mixed model, but modified so states were required to select a certain number of optional components.

This response indicates that state-by-state assessment of student achievement should be based upon specification in each subject area that represents the broadest statement of instructional content and priorities that the states can agree upon. It further suggests that the development of the assessment program can proceed, with the convening of subject-matter and assessment specialists from the states to discuss and establish these subject-area frameworks.

Information to be Provided

While plans for the assessment program included scaling, no recommendations were made regarding the type and level of information that the assessment program would provide. States were asked to comment on three issues: what level of detail should the assessment program provide in reporting results; specifically, should only a single, global value be reported for performance in each subject at each grade level, or should more detailed information be included on performance in major sub-areas in each subject--comprehension, vocabulary, and word attack in reading, or computation and problem-solving in mathematics? Second, how frequently should performance in the various subjects be assessed and reported; i.e. yearly? Every two years? Every four years? Finally, what type of scale should be used to report results; specifically, should the scales used in the current National Assessment of Educational Progress be used, or should some other scale be developed and used?

On the issue of the level of detail to be provided, states were consistently in favor of reporting simple, global information on performance in each subject. Seventeen states preferred global scores and just two preferred that the information be reported only in terms of sub-areas. Eight states indicated that subscales could be reported usefully as options or as secondary information in reports.

This suggests that information from a state-by-state assessment program be reported as single values for each subject at each grade level, perhaps offering more detailed information on instructional sub-areas or as optional or additional information.

On the issue of the type of scale to be used, participants in the March meeting were split between preferring to use a scale consistent with the National Assessment of Educational Progress and preferring to develop a unique scale for a new, state-by-state assessment program. Few states indicated a preference on this issue in response to the planning paper, suggesting that this decision be deferred until it can be determined whether the National Assessment item pool and scales can be used in the state-by-state assessment program.

On the issue of the schedule and cycle with which to collect and report assessment information in each subject, states were presented with four options ranging from assessing all five subjects (reading, mathematics, science, social studies and English) every year to assessing one subject each year and each subject every four years (if reading and English were combined). The following figure shows these options.

Responses by states to these options were mixed. No state indicated that it was most desirable to assess every subject each year, but opinions were divided among the other options. Six states indicated that relatively infrequent assessment--one subject each year through a four-year cycle--was most desirable. Four states argued for the cycle used currently in the National Assessment of Educational Progress: assessing reading every two years along with either science and mathematics or writing and social studies, so reading results would be available every two years and results in the other subjects would be provided every four years. Eight states opted for the other intermediate cycle: assessing reading, writing and English in combination one year and mathematics the next, alternating science and social studies with mathematics. This would produce results for the basic skills areas every two years and the other subjects every four years. The difference in cost for the two intermediate options was estimated at about $50-55,000 on a yearly basis, with the NAEP schedule being lower in cost.

Figure

OPTIONAL SCHEDULES FOR SUBJECT-AREA ASSESSMENT

Year	Option 1 One Subject Each Year	Option 2 NAEP Schedule (Reading Plus One Other Area Every 2 Years)	Option 3 Basic Skills Alternating, Other Areas Every 4 Years	Option 4 Every Subject, Every Year
1986		Rdg, Math Sci (NAEP)		
1987				
1988		Rdg, Wrtg SS (NAEP)		
- - - - - - - - -	CCSSO Assessment Program Begins			- - - - - - -
1989	Rdg/Wrtg/Eng	Rdg, Math Sci	Rdg, Wrtg/Eng	Rdg, Wrtg/Eng Math, Sci, SS
1990	Math		Math, Sci	Rdg, Wrtg/Eng Math, Sci, SS
1991	Sci	Rdg, Wrtg SS	Rdg, Wrtg/Eng	Rdg, Wrtg/Eng Math, Sci, SS
1992	SS		Math, SS	Rdg, Wrtg/Eng Math, Sci, SS
1993	Rdg/Wrtg/Eng	Rdg, Math Sci	Rdg, Wrtg/Eng	.
1994	.	.	Math, Sci	.
	.	.	.	
	.	.	.	
			.	

Key:
Rdg = Reading
Math = Mathematics
Sci = Science
Wrtg = Writing
SS = Social Studies
Rdg/Wrtg/Eng = Combined Reading, Writing, Literature and Language Arts
Wrtg/Eng = Combined Writing and English Literature and Language Arts

Year = Year of data collection and reporting

NAEP = Current NAEP program

These results clearly suggest that an intermediate schedule is desired by states. On the basis of the numbers of responses, Option 3, or basic skills plus the other subjects, is felt marginally to be better.

Standardizing Procedures

Finally, states were asked how they felt procedures for standardizing data collection across states should be established. All respondents indicated that procedures should be governed by explicit guidelines established by CCSSO or another suitable organization, that sampling and test-administration should be monitored to ensure uniformity, and that sampling, scoring, and data-processing should be conducted through a single, central project funded through pooling of resources from the states and administered by CCSSO.